Cyclostationarity-based Robust Algorithms for QAM Signal Identification

Octavia A. Dobre, *Senior Member, IEEE*, Menguc Oner, *Member, IEEE*, Sreeraman Rajan, *Senior Member, IEEE*, and Robert Inkol, *Senior Member, IEEE*

***Abstract*—This letter proposes two novel algorithms for the identification of quadrature amplitude modulation (QAM) signals. The cyclostationarity-based features used by these algorithms are robust with respect to timing, phase, and frequency offsets, and phase noise. Based on theoretical analysis and simulations, the identification performance of the proposed algorithms compares favorably with that of alternative approaches.**

***Index Terms*—Quadrature amplitude modulation, Signal cyclostationarity, Signal identification.**

I. Introduction

SIGNAL identification, a major task of intelligent receivers, finds applications in software defined and cognitive radios, and spectrum surveillance and management. Signal identification algorithms can be grouped into two general classes: likelihood-based and feature-based (see the comprehensive survey provided in [1] and references therein). The former tends to be complex to implement and sensitive to model mismatches, such as timing, phase, and frequency offsets, and phase noise [1]–[4]. The latter can be simple to implement and be made robust to some model mismatches through the careful selection of features [1], [5]–[9]. This letter proposes two novel algorithms for the identification of quadrature amplitude modulation (QAM) signals. The cyclostationarity-based features used by these algorithms are robust with respect to timing, phase, and frequency offsets, and phase noise. The rest of the letter is organized as follows. Signal model and its statistical characterization are presented in Sections II and III, respectively. Feature selection is explained in Section III, and the proposed algorithms are described in Section IV. A comparative performance analysis is presented in Section V, followed by final remarks in Section VI.

II. Signal Model

The baseband received waveform is expressed as

$$r(t) = \alpha e^{j(2\pi \Delta f_c t + \theta + \varphi(t))} \sum_k s_k^{(i)} p(t - kT - \epsilon T) + w(t), \quad (1)$$

where α is the signal amplitude, Δf_c is the frequency offset, θ is the phase offset, $\varphi(t)$ is the phase noise (PN), $s_k^{(i)}$ is the symbol transmitted within the kth period corresponding to an $\Omega^{(i)}$ QAM modulation, $p(t)$ is the overall pulse shape, T is the symbol period, ϵ is the timing offset, and $w(t)$ is additive zero-mean complex Gaussian noise. The symbol sequence $\{s_k^{(i)}\}$ is a zero-mean independent and identically distributed sequence, with values drawn from an alphabet corresponding to the $\Omega^{(i)}$ signal constellation. Without loss of generality, we consider unit variance constellations, i.e., $c_{s,2,1}^{(i)} = E\{|s_k^{(i)}|^2\} = 1$, with $E\{.\}$ as the expectation operator. The PN $\varphi(t)$ is modeled (asymptotically with time) as a Wiener random process with mean zero and variance $2\pi B_L|\, t\,|$, where B_L is the full 3 dB bandwidth of the Lorentzian power spectrum [10].

III. Statistical Signal Characterization and Feature Selection

In the absence of the PN, the received signal $r(t)$ is nth-order cyclostationary, with the (n, q) (nth-order/ q-conjugate) cyclic cumulants (CCs) and cycle frequencies (CFs) given respectively as [5]–[7] [1]

$$\begin{aligned} c_r(\gamma_k;\tau)_{n,q} &= \alpha^n c_{s,n,q}^{(i)} T^{-1} e^{-j2\pi\beta_k \epsilon T} e^{j\theta(n-2q)} \\ &\times e^{j2\pi\Delta f_c \sum_{m=1}^{n} (-)_m \tau_m} \int_{-\infty}^{\infty} \prod_{m=1}^{n} p^{(*)_m}(t+\tau_m) \\ &\times e^{-j2\pi\beta_k t} dt + \, c_w(\gamma_k;\tau)_{n,q}, \end{aligned} \quad (2)$$

$$\gamma_k = \beta_k + (n-2q)\Delta f_c, \beta_k = kT^{-1}, k \text{ integer}, \quad (3)$$

where $\tau = [\tau_{\mathbf{1}}, ..., \tau_{\mathbf{n-1}}, \tau_{\mathbf{n}} = 0]$ is the delay-vector, $c_{s,n,q}^{(i)}$ is the (n, q) cumulant of the $\Omega^{(i)}$ signal constellation, $(*)_m$ is the optional conjugation of the mth term so that the number of conjugations equals q, $(-)_m$ is the minus sign associated to the optional conjugation $(*)_m, m = 1, ..., n$, and $c_w(\gamma_k;\tau)_{n,q}$ is the (n, q) CC of $w(t)$ at CF γ_k and delay vector τ. For examples of $c_{s,n,q}^{(i)}$, $\Omega^{(i)}$ =4-QAM, 16-QAM, 64-QAM, and V.29 [2], and diverse n and q values, see Table I [5]–[9]. Note that $c_{s,n,q}^{(i)} = c_{s,n,n-q}^{(i)}$ for n even and $c_{s,n,q}^{(i)} = 0$ for n odd, $q = 0, ..., n$.

TABLE I
Cumulant Values for QAM Signal Constellations

	$c_{s,2,0}$	$c_{s,2,1}$	$c_{s,4,1}$	$c_{s,4,2}$	$c_{s,6,3}$
4-QAM	0	1	0	-1	4
16-QAM	0	1	0	-0.68	2.08
64-QAM	0	1	0	-0.619	1.797
V.29	0	1	0	-0.5816	1.4897

Manuscript received September 25, 2011. The Associate Editor coordinating the review of this letter was Georgios Ginis.

This work was supported in part by the Defence Research and Development Canada, Ottawa.

O. A. Dobre is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, Canada. E-mail: odobre@mun.ca.

M. Oner is with the Department of Electronics Engineering, Isik University, Istanbul, Turkey. E-mail: oner@isikun.edu.tr.

S. Rajan and R. Inkol are with the Defence Research and Development, Ottawa, Canada. E-mail: {sreeraman.rajan, robert.inkol}@drdc-rddc.gc.ca.

Digital Object Identifier

[1] For the definitions of the nth-order cyclostationarity, corresponding time-varying and cyclic statistics, and their estimators, the reader is referred to [11], [12].

By inspection of (2) and (3), for certain values of the parameters k, q, and τ, specifically $k = 0$, $q = n/2$, [2] and $\sum_{m=1}^{n}(-)_m\tau_m = 0$, the corresponding CCs are not dependent on ϵ, θ, and Δf_c, respectively. Furthermore, from (2) and Table I, it is clear that the second-order CCs cannot serve as discriminating features for QAMs, and we need to resort to higher-order statistics, e.g., $n = 4, 6$. In such a case, $c_w(\gamma_k;\tau)_{n,q}$ is zero for any γ_k and τ, due to the Gaussian nature of the noise.

Moreover, one can show that choosing $q = n/2$ and zero delay-vector does not only cancel the effect of θ and Δf_c, respectively, but also the effect of the PN. In this case, the (n, q) CCs, n =4, 6, are practically given by (2). A sketch of the proof is as follows: - The presence of the PN leads to the appearance of the multiplicative term $e^{-(n-2q)^2\pi B_L t}$ in the expression for the (n, q) time-varying moments of the signal component at zero delay vector [13]. Clearly, for $q = n/2$, the effect of this multiplicative term vanishes, and the PN has no effect on these moments; - The (4,2) and (6,3) time-varying cumulants at zero delay-vector are respectively expressed as functions of (4,2) and (6,3) and lower-order time-varying moments through the moment-to-cumulant formula [11]; - By employing results from the first step, the fact that the (2,0) and (2,2) time-varying moments and cumulants are equal, the (2,0) and (2,2) time-varying cumulants are a function of CCs [11], and the (2,0) and (2,2) cumulants of the QAM signal constellations vanish, one can further show that the PN has no effect on the (4,2) and (6,3) time-varying cumulants of QAM signals at zero delay-vector; - By using the CC expression as a function of the time-varying cumulant [11], one can finally conclude that the (4,2) and (6,3) CCs at zero delay-vector are not affected by the PN.

Hence, we choose the (4,2) and (6,3) CCs at zero CF and zero delay-vector as discriminating features, due to their robustness to ϵ, θ, Δf_c, and $\varphi(t)$.

IV. Proposed Algorithms

We formulate QAM signal identification as a multiple hypothesis testing problem, i.e., under hypothesis H_i we decide that the modulation is $\Omega^{(i)}$, $i = 1, ..., N_H$. Here N_H represents the number of hypotheses, which are equally likely. We propose two identification algorithms, as follows:

Algorithm I: An estimate of the feature vector is obtained from data, $\hat{\mathbf{F}}_{\mathrm{I}} = [\hat{c}_r(0;0)_{4,2}, \hat{c}_r(0;0)_{6,3}]$, and used with an optimum criterion to make a decision. Based on the asymptotic bivariate Gaussian distribution of this estimate [12], the optimum decision criterion in the sense that it maximizes the average probability of correct identification, is

$$i = \mathrm{argmin}_{i=1,...,N_H}(\hat{\mathbf{F}}_{\mathrm{I}} - \mathbf{F}_{\mathrm{I}}^{(i)})\Sigma^{-1}(\hat{\mathbf{F}}_{\mathrm{I}} - \mathbf{F}_{\mathrm{I}}^{(i)})^T, \quad (4)$$

where $\mathbf{F}_{\mathrm{I}}^{(i)}$ represents the mean of $\hat{\mathbf{F}}_{\mathrm{I}}$ under hypothesis H_i, Σ is the covariance matrix of $\hat{\mathbf{F}}_{\mathrm{I}}$, and the superscripts $^{-1}$ and ^{T} represent the inverse and transpose operations, respectively. Note that equal covariance matrices are considered under all hypotheses, as these depend weakly on the modulation, and strongly on the observation interval, signal-to-noise ratio (SNR), CC order, and CF [6]. Moreover, when applying (4), we use an estimate of the covariance matrix, $\hat{\Sigma}$, which is calculated based on the observed data. For these estimators, the reader is referred to eq. (118) in [12]; we should mention that the covariance estimators are expressed as summations of products of moment and smoothed cross cyclic periodogram estimators.

Algorithm II: An estimate of the feature vector is obtained from data, $\hat{\mathbf{F}}_{\mathrm{II}} = [\hat{c}_r^{1/2}(0;0)_{4,2}, \hat{c}_r^{1/3}(0;0)_{6,3}]$, and used to make a decision

$$i = \mathrm{argmin}_{i=1,...,N_H}\|\hat{\mathbf{F}}_{\mathrm{II}} - \mathbf{F}_{\mathrm{II}}^{(i)}\|_2, \quad (5)$$

where $\|.\|_2$ represents the vector 2-norm, and $\mathbf{F}_{\mathrm{II}}^{(i)}$ is the mean of $\hat{\mathbf{F}}_{\mathrm{II}}$ under hypothesis H_i.

When compared with Algorithm I, the features consist of the (n, q) CCs, $n = 4, 6$, and $q = n/2$, raised to the power of $2/n$. This operation renders closer variances of the feature estimates [6], and we adopt the minimum 2-norm as the decision-making criterion. Clearly, Algorithm II is simpler than Algorithm I, as it does not require estimation of the covariance matrix.

The proposed algorithms are formally stated below. Note that they do not require estimation and compensation for the timing, phase, and frequency offsets, and phase noise.

Proposed Algorithms

1: **INPUT** Pool of QAM modulations to be classified, $\{\Omega^{(i)}\}$, and corresponding $c_{s,4,2}^{(i)}$, $c_{s,6,3}^{(i)}$, $i = 1, ..., N_H$.
2: Estimate the signal bandwidth [14], and coarsely estimate the carrier frequency as the middle point of the estimated bandwidth. Remove the out-of-band noise by filtering, and down-convert the signal.
3: Oversample the signal and estimate the corresponding feature vector from data.
4: Calculate $\mathbf{F}_{\mathrm{I}}^{(i)}$ for Algorithm I, or $\mathbf{F}_{\mathrm{II}}^{(i)}$ for Algorithm II, $i = 1, ..., N_H$. Based on the fact that the CC estimators are asymptotically unbiased [11], [12], (2) is used to obtain the mean of each feature estimate.
5: Make a decision by applying the corresponding criterion:
   - For Algorithm I, estimate the covariance matrix from data, and apply the criterion in (4).
   - For Algorithm II, directly apply (5).
6: **OUTPUT** Selected hypothesis, H_i, and corresponding QAM modulation, $\Omega^{(i)}$.

V. Performance Evaluation

The performance of Algorithm I is investigated through theoretical analysis and simulations. Additional simulations are employed to compare the performance of Algorithms I and II with the algorithms proposed in [3]–[9].

A. Theoretical Performance Analysis: Algorithm I

Following [15], [16], the probability of correct identification for the case of a binary classification problem ($N_H = 2$) is

[2]Note that $k = 0$ and $q = n/2$ render CCs at zero CF; however, these are different from "stationary" cumulants, as their calculation involves products of lower-order cyclic moments at non-zero CFs (see eq. (42) in [11]).

given by

$$\begin{aligned} P_c &= 2^{-1}(P(H_1|H_1) + P(H_2|H_2)) \\ &= 2^{-1}\left(\int_{-\infty}^{-\mu^{(1)}/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du + \int_{-\mu^{(2)}/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du\right) \\ &= Q(-\sqrt{\Delta\mathbf{F}\Sigma^{-1}\Delta\mathbf{F}^T}/2), \end{aligned} \tag{6}$$

where $P(H_i|H_i)$, $i = 1, 2$, represents the probability to decide for the hypothesis H_i when indeed this is true, $\mu^{(1)} = -\Delta\mathbf{F}\Sigma^{-1}\Delta\mathbf{F}^T$, $\mu^{(2)} = \Delta\mathbf{F}\Sigma^{-1}\Delta\mathbf{F}^T$, $\sigma^2 = 4\Delta\mathbf{F}\Sigma^{-1}\Delta\mathbf{F}^T$, with $\Delta\mathbf{F} = \mathbf{F}_{\mathrm{I}}^{(2)} - \mathbf{F}_{\mathrm{I}}^{(1)}$, and $Q(.)$ is the Q function, defined as $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$. For a multiple hypothesis testing problem, a general closed-form expression for the probability of correct identification is not available, and one usually resorts to performance bounds [16], [17]. The union bound represents a commonly used upper bound for the probability of error, while its pruned versions can yield closer upper bounds or approximations of this probability [17]. These in turn lead to lower bounds or approximations for the probability of correct identification.

B. Simulation and Numerical Results

In order to illustrate the performance of the proposed algorithms, we consider two modulation pools, $\eta_1 = \{$4-QAM, 16-QAM $\}$ ($N_H = 2$), and $\eta_2 = \{$4-QAM, 16-QAM, V.29$\}$($N_H = 3$). A root raised cosine pulse shape with 0.35 roll-off factor is employed at the transmit- and receive-sides. At the receiver, the signal is oversampled by a factor of 11. The model mismatches are set as follows: the phase offset θ is uniformly distributed over $[-\pi, \pi)$, the timing offset ϵ is uniformly distributed over $[0, 1)$, and the normalized carrier frequency offset and phase noise bandwidth are respectively equal to $\Delta f_c T = 10^{-2}$ and $B_L T = 10^{-3}$, unless otherwise mentioned. The in-band SNR is considered, and $\alpha = 1$ unless otherwise mentioned. The number of trials used to calculate $P(H_i|H_i), i = 1, ..., N_H$, is 10^3.

Fig. 1 plots the average probability of correct identification P_c (for the modulation pool η_1) versus SNR, when using 1,000 and 2,000 symbols for estimation, respectively. The simulation and theoretical results for Algorithm I are in agreement. Note that the performance of Algorithms I and II converges as the SNR increases. For higher values of P_c ($P_c \geq 0.9$), which are of most practical interest, the two algorithms compare closely in performance. As expected, a better performance is achieved with a larger observation interval.

For the same model mismatches, number of symbols, and SNRs, simulations demonstrated that the algorithms developed in [3], [4], [6]–[9] fail (P_c below 0.6) for diverse reasons: the algorithm in [3] is affected by timing and frequency offsets, and phase noise, the algorithms in [4], [6], [7] are affected by frequency offset and phase noise, the algorithm in [8] is affected by timing offset, and the one in [9] is affected by timing and frequency offsets, as well as phase noise. The effectiveness of the proposed algorithms is further confirmed by additional performance comparisons with the algorithm in [6]. While the algorithm in [6] requires $\Delta f_c T = 10^{-5}$ and $B_L T = 10^{-6}$ to achieve a probability of correct classification approaching 1 for 1,000 symbols and 10 dB SNR, the proposed algorithms achieve such a performance for $\Delta f_c T = 10^{-2}$ and $B_L T = 10^{-3}$. These results clearly show the effectiveness of the proposed algorithms.

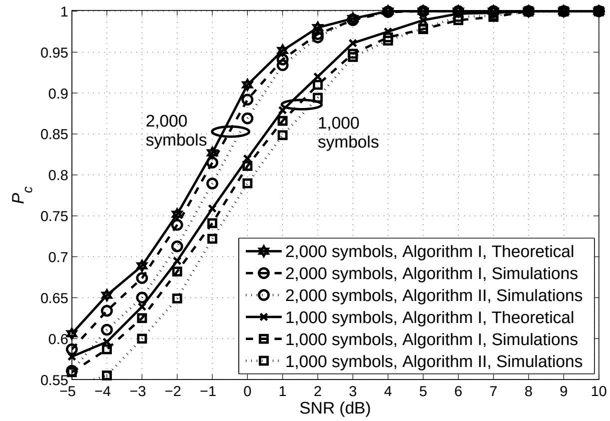

Fig. 1. Performance of the proposed algorithms for the modulation pool η_1.

Fig. 2 shows the average probability of correct identification P_c versus SNR achieved with Algorithms I and II for the modulation pool η_2, when using 3,000 and 5,000 symbols for estimation, respectively. A lower bound on the performance of Algorithm I is also depicted, which is obtained by truncating the union bound. Based on the fact that the constellation points are approximately co-linear in the feature space, we retain those terms in the union bound which correspond to such a particular geometry. Note that this yields a tight lower bound for the average probability of classification of Algorithm I at lower SNR, while it provides a good approximation at higher SNR. Furthermore, the performance of Algorithms I and II converges as the SNR increases. As it was observed for η_1, the two algorithms compare closely in performance for $P_c \geq 0.9$, which is of most practical interest. When compared to the results obtained for η_1, a larger number of symbols is needed to achieve a P_c approaching 1 for the same SNR. For example, 3,800 symbols were required with Algorithm II to approach a P_c of 1 at 10 dB SNR. A P_c of 0.982, 0.955, and 0.858 was respectively achieved for 3,000, 2,000, and 1,000 symbols and 10 dB SNR. It is worth noting that 4-QAM was perfectly identified in all cases, and the performance degradation resulted from the mis-identification of 16-QAM and V.29. Moreover, it was observed that, regardless of the SNR, the ability of P_c to approach 1 was adversely affected when the number of symbols was reduced. This phenomenon can be explained as follows: the distance between 16-QAM and V.29 in the feature space is smaller, and thus, the correct decision between these two modulation types is particularly sensitive to the accuracy of the feature estimates. Clearly, for the observation intervals of 1,000 and 2,000 symbols, the effect of the estimation errors dominates and leads to an irreducible error floor, irrespective of the SNR.

The proposed algorithms are also applicable under block fading channel conditions. For example, a P_c of 0.962 is obtained for 1,000 symbols and 10 dB (average) SNR with Algorithm II for the modulation pool η_1 ($N_H = 2$). This performance can be further enhanced by exploiting spatial receive diversity [7]. For example, using two receive antennas and selection combining, a P_c approaching 1 is achieved under the above conditions.

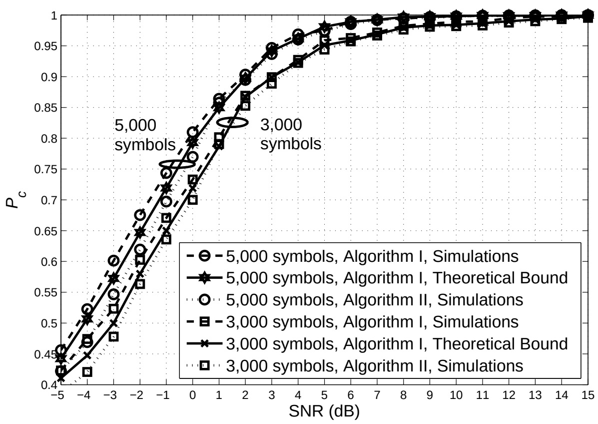

Fig. 2. Performance of the proposed algorithms for the modulation pool η_2.

VI. Conclusion

Two cyclostationarity-based algorithms have been proposed and evaluated for the identification of QAM signals. These algorithms are robust with respect to timing, phase, and frequency offsets, and phase noise. Furthermore, the version using the minimum 2-norm criterion has the advantage of implementation simplicity. Simulation results demonstrate that they compare favorably with other algorithms described in the literature.

Ackowledgment

The authors are grateful to the anonymous reviewers and the Associate Editor, Dr. G. Ginis, for their constructive comments on the paper.

References

[1] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, ”A survey of automatic modulation classification techniques: classical approaches and new developments,” *IET Comm.*, vol. 1, pp. 137-156, Apr. 2007.
[2] W. Wei and J. M. Mendel, ”Maximum likelihood classification for digital amplitude phase modulations,” *IEEE Trans.Commun.*, vol. 42, pp. 189-193, Feb. 2000.
[3] F. Hameed, O. A. Dobre, and D. C. Popescu, ”On the likelihood-based approach to modulation classification,” *IEEE Trans. Wireless Commun.*, vol. 8. pp. 5884-5892, Dec. 2009.
[4] W. C. Headley and C. R. C. M. da Silva, ”Asynchronous classification of digital amplitude-phase modulated signals in flat fading channels,” *IEEE Trans. Commun.*, vol. 59, pp. 7-12, Jan. 2011.
[5] C. M. Spooner, W. A. Brown, and G. K. Yeung, ”Automatic radio-frequency environment analysis,” *IEEE ASILOMAR*, 2000, pp. 1181-1186.
[6] C. M. Spooner, ”On the utility of sixth-order cyclic cumulants for RF signal classification,” *IEEE ASILOMAR*, 2001, pp. 890-897.
[7] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, ”Cyclostationarity-based modulation classification of linear digital modulations in flat fading channels,” *Kluwer Wireless Personal Communications Journal*, vol. 54, pp. 699-720, Sep. 2010.
[8] H.-C. Wu, M. Saquib, and Z. Yun, ”Novel automatic modulation classification using cumulant features for communications via multipath channels,” *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3098-3105, Aug. 2008.
[9] A. Swami and B. M. Sadler, ”Hierarchical digital modulation classification using cumulants,” *IEEE Trans. Commun.*, vol. 48, pp. 416-429, Mar. 2000.
[10] A. Demir, A. Mehrotra, and J. Roychowdhury, ”Phase noise in oscillators: a unifying theory and numerical methods for characterization,” *IEEE Trans. Circ. Syst.- Part I: Fundamental Theory and Applicat.*, vol. 47, pp. 655-674, May 2000.
[11] W. A. Gardner and C. M. Spooner, ”The cumulant theory of cyclostationary time-series, Part I: Foundation,” *IEEE Trans. Sig. Proc.*, vol. 42, 3387-3408, Dec. 1994.
[12] A. V. Dandawate and G. B. Giannakis, ”Asymptotic theory of mixed time averages and *k*th-order cyclic-moment and cumulant statistics,” *IEEE Trans. Inf. Th.*, vol. 41, pp. 216-232, Jan. 1995.
[13] O. A. Dobre, Y. Bar-Ness, and W. Su, ”Robust QAM modulation classification using cyclic cumulants,” *IEEE WCNC*, 2004, pp. 745-748.
[14] Q. Chen, Y. Wang, C. W. Bostian, ”Universal classifier synchronizer demodulator,” *IEEE IPCCC*, 2008, pp. 366-371.
[15] K. Fukunaga and T. Krile, ”Calculation of Bayes’ recognition error for two multivariate Gaussian distributions,” *IEEE Trans. Computers*, vol. C-18, pp. 220-229, Mar. 1969.
[16] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. 2nd ed., John Wiley and Sons, 2001.
[17] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.